ỨNG DỤNG THUẬT TOÁN FACENET XÂY DỰNG HỆ THỐNG NHẬN DẠNG KHUÔN MẶT

APPLYING THE FACENET ALGORITHM TO DEVELOP FACE RECOGNITION SYSTEM

Mai Văn Hà[1*], Nguyễn Thế Xuân Ly[1]

[1]Trường Đại học Bách khoa - Đại học Đà Nẵng

*Tác giả liên hệ: mvha@dut.udn.vn

(Nhận bài: 18/6/2021; Chấp nhận đăng: 15/7/2021)

Tóm tắt - Nhận dạng khuôn mặt là một bài toán phổ biến đang đặt ra hiện nay. Tồn tại nhiều phương pháp và hướng tiếp cận đối với bài toán nhận dạng khuôn mặt: Tiếp cận theo đặc trưng toàn cục (sử dụng các đặc điểm toàn cục của khuôn mặt) và tiếp cận theo đặc trưng cục bộ (sử dụng các đặc điểm cục bộ của khuôn mặt). Tuy nhiên hiệu quả của các phương pháp nhận dạng này vẫn còn hạn chế và độ chính xác chưa cao khi dữ liệu đầu vào bị ảnh hưởng bởi các yếu tố khách quan của môi trường (độ sáng, hướng nghiêng, kích thước, …). Do đó, nhóm tác giả đề xuất xây dựng hệ thống nhận dạng khuôn mặt dựa trên thuật toán FaceNet và sử dụng Multi-task Cascaded Convolutional Networks phát hiện và xác định khuôn mặt cho phép nâng cao hiệu quả nhận dạng.

Từ khóa - Nhận dạng khuôn mặt; phát hiện khuôn mặt; thuật toán FaceNet; Multi-task Cascaded Convolutional Networks (MTCNN)

Abstract - Face recognition is a popular problem being mentioned these days. There are some methods and approaches to deal with this problem: the global one (using global features of the face) and local one (applying local features of the face). However, the effectiveness of those identification methods is still limited and the accuracy is not high when the input data is affected by the objective factors of environment such as brightness, tilt direction, size and so on. Therefore, the authors propose developing a face recognition system based on the Face Net algorithm and the use of Multi-task Cascaded Convolutional Networks while detecting as well as identifying faces in images to improve the recognition efficiency.

Key words - Face identification; Face verification; FaceNet algorithm; Multi-task Cascaded Convolutional Networks (MTCNN)

1. Đặt vấn đề

Cùng với sự phát triển của xã hội, vấn đề an ninh bảo mật là một điều tất yếu hiện nay. Các hệ thống nhận dạng con người được ra đời với độ tin cậy ngày càng cao. Có thể kể đến như nhận dạng hình dáng, nhận dạng giọng nói, nhận dạng khuôn mặt,... Trong đó, phổ biến và được ứng dụng nhiều hơn cả là bài toán nhận dạng khuôn mặt.

Hiện nay, tồn tại một số hướng tiếp cận đối với bài toán nhận dạng khuôn mặt: Tiếp cận theo đặc trưng toàn cục và tiếp cận theo đặc trưng cục bộ.

Đối với phương pháp tiếp cận theo hướng toàn cục thì các đặc trưng chung của khuôn mặt sẽ được sử dụng để nhận dạng như: Màu sắc, hình dạng, các nét chính của khuôn mặt… Phương pháp được sử dụng phổ biến trong hướng tiếp cận này là Eigengaces-PCA và Fisherfaces.

Phương pháp Eigenfaces sử dụng phép phân tích thành phần (Principal Components Analysis – PCA) cho phép giảm số chiều dữ liệu. Với phương pháp này, sau quá trình chuẩn hoá, các đặc trưng toàn cục của khuôn mặt sẽ được biểu diễn thành các véc-tơ riêng. Tập hợp các véc-tơ này tạo thành không gian mới với số chiều dữ liệu giảm xuống mà các đặc trưng quan trọng của khuôn mặt vẫn được giữ lại trong quá trình nhận dạng. Trong không gian véc-tơ này, mỗi véc-tơ được gọi là Eigenfaces. Do PCA là thuật toán học không có giám sát nên có hạn chế trong trường hợp tập dữ liệu huấn luyện có nhiều hơn một mẫu cho mỗi lớp.

Phương pháp Fisherfaces là phương pháp phân tích tuyến tính khác biệt (Linear Discriminant Analysis – LDA) đã được sử dụng nhằm khai thác tốt hơn các thông tin về lớp nói trên [2]. LDA cho phép nhận diện khuôn mặt dựa trên một phép chiếu tuyến tính từ không gian hình ảnh vào một chiều không gian thấp hơn bằng cách tối đa giữa các lớp tán xạ và giảm thiểu phân tán trong lớp. Có thể thấy rằng, phương pháp LDA áp dụng các tiêu chuẩn phân biệt tuyến tính cho phép tối đa hóa tỷ lệ yếu tố quyết định của lớp giữa ma trận tán xạ của các lớp do đó cho phép khắc phục những nhược điểm của phương pháp Eigengaces-PCA.

Hướng tiếp cận nhận dạng dựa trên các đặc trưng cục bộ của khuôn mặt như: Các chi tiết như mắt, mũi, lông mày, điểm ảnh, … Hướng tiếp cận này sử dụng hai phương pháp phổ biến là phương pháp lấy mẫu nhị phân cục bộ (Local Binary Pattern – LBP) và phương pháp biến đổi sóng nhỏ Gabor (Gabor wavelets) [3]. Trong LBP, bức ảnh sẽ được chia thành các vùng bằng nhau, tại mỗi vùng này có thể tính được 1 LBP histogram và dựa vào đó xác định được thông tin về vị trí mắt, mũi, miệng trên khuôn mặt. Các thông tin này áp dụng trọng số lên histogram của các vùng chứa các đặc trưng quan trọng cho phép phân biệt giữa các khuôn mặt. Với phương pháp Gabor wavelets thì dữ liệu được chia thành các thành phần với tần số khác nhau và xem xét từng thành phần với độ phân giải thích hợp [4]. Với phương pháp này các ảnh khuôn mặt sẽ được trích chọn đặc trưng dựa vào biến đổi Gabor wavelet. Một tập các tần số và hướng của các điểm đặc trưng xác định bởi mạng wavelet sẽ là thông tin đặc trưng để biểu diễn ảnh.

Tuy nhiên, việc xây dựng hệ thống nhận dạng khuôn mặt với hiệu suất và độ chính xác cao là một thách thức rất

[1] The University of Danang - University of Science and Technology (Mai Van Ha, Nguyen The Xuan Ly)

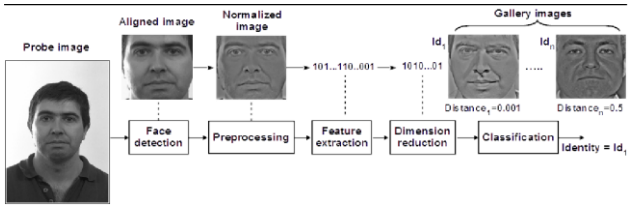
lớn vì những yếu tố khách quan như môi trường, ánh sáng, độ nghiêng của khuôn mặt, độ tuổi, cảm xúc hay như việc bị che khuất. Vì vậy, việc xây dựng một hệ thống nhận dạng khuôn mặt hoạt động tốt dù khuôn mặt bị che lấp một phần hay bị ảnh hưởng bởi các yếu tố của môi trường xung quanh là cần thiết.

Do đó, nhóm tác giả đề xuất sử dụng thuật toán FaceNet để nhận dạng khuôn mặt và ứng dụng Multi-task Cascaded Convolutional Networks (MTCNN) đối với việc phát hiện khuôn mặt trong bức ảnh.

2. Ứng dụng thuật toán FaceNet trong nhận dạng khuôn mặt

2.1. Tổng quan bài toán nhận dạng khuôn mặt

Nhận dạng khuôn mặt người là một chủ đề nghiên cứu thuộc lĩnh vực thị giác máy được phát triển từ nhưng năm 90 của thế kỷ trước. Hiện nay, lĩnh vực nhận dạng được đẩy mạnh phát triển và nhận được sự quan tâm của nhiều nhà nghiên cứu từ nhiều lĩnh vực nghiên cứu khác nhau đặc biệt là nhận dạng khuôn mặt.



Hình 1. *Mô hình chung của bài toán nhận dạng khuôn mặt*

Bài toán nhận dạng khuôn mặt hướng tiếp cận cũng tương tự như hệ thống thị giác của con người khi cần nhận dạng một ai đó khi nhìn vào 1 bức ảnh. Hoạt động của hệ thống nhận dạng khuôn mặt có được triển khai chi tiết như sau:

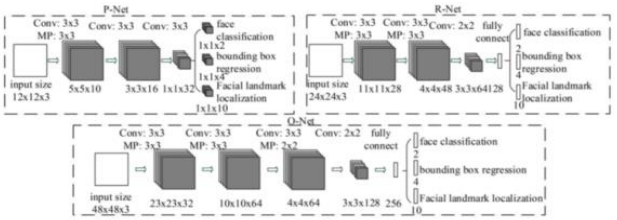
- Bước 1: Phát hiện và xác định khuôn mặt trong bức ảnh.

- Bước 2: Chuẩn hoá và trích chọn đặc trưng khuôn mặt đã được phát hiện trong bước 1.

- Bước 3: Tiến hành so sánh và nhận dạng các đặc trưng ở bước 2 với tập dữ liệu huấn luyện đã có để đưa ra kết quả kết luận nhận dạng.

2.2. Sử dụng MTCNN phát hiện khuôn mặt

Vấn đề đầu tiên của nhận dạng khuôn mặt là phải phát hiện và xác định được vị trí khuôn mặt trọng bức ảnh. Trong bài báo này nhóm tác giả đề xuất sử dụng MTCNN để phát hiện và xác định khuôn mặt người trong bức ảnh [5].

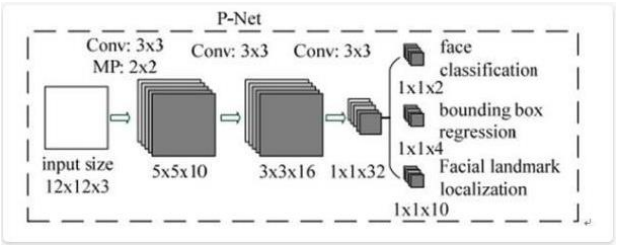


Hình 2. *Sơ đồ hoạt động của MTCNN [5]*

Về mặt cấu trúc MTCNN bao gồm 3 mạng CNN (Convolutional Neural Networks) xếp chồng và đồng thời hoạt động khi phát hiện và xác định khuôn mặt. Mỗi mạng CNN trong MTCNN có cấu trúc và vai trò khác nhau trong việc phát hiện khuôn mặt. Kết quả dữ liệu đầu ra của MTCNN là véc-tơ đặc trưng biểu diễn cho vị trí khuôn mặt được xác định trong bức ảnh (mắt, mũi, miệng, …)

MTCNN hoạt động theo 3 bước với 3 mạng nơ-ron riêng cho mỗi bước (P-Net, R-Net và O-Net). Khi sử dụng, MTCNN sẽ cho phép tạo ra nhiều bản sao của hình ảnh đầu vào, với các kích thước khác nhau để làm dữ liệu đầu vào.

Tầng 1: Sử dụng mạng CNN, gọi là Mạng đề xuất (P-Net), để thu được các cửa sổ chứa khuôn mặt và các vectơ hồi quy trong các cửa sổ đó. Tiếp theo, các cửa sổ chứa khuôn mặt được hiệu chuẩn dựa trên các vector hồi quy. Cuối cùng, những cửa sổ xếp chồng nhau tại một vùng được hợp nhất thành một cửa sổ. Kết quả đầu ra là các cửa sổ có thể chứa khuôn mặt.



Hình 3. *Mạng đề xuất (P-Net) [5]*

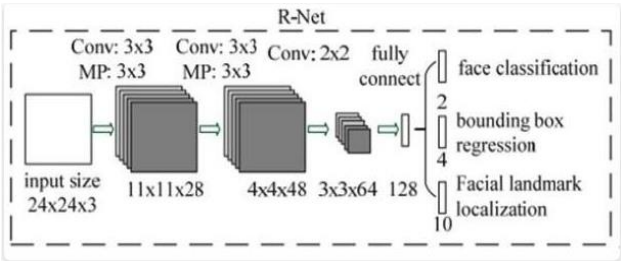
Mạng P-Net sử dụng kiến trúc CNN gồm 3 lớp tích chập và 1 lớp co. Đầu vào cửa sổ trượt với kích thước 12x12x3 (với 3 tương ứng với 3 màu: Đỏ, xanh lục, xanh lam trong hệ màu RGB thông thường). Kết quả của P-Net gồm 3 cụm như sau:

- Cụm thứ nhất có 2 bộ lọc kích thước 1x1 nhận dạng khuôn mặt;

- Cụm thứ hai có 4 bộ lọc kích thước 1x1 đóng khung 4 vị trí hộp giới hạn;

- Cụm thứ ba có 10 bộ lọc kích thước 1x1 đóng khung 10 vị trí khuôn mặt.

Tầng 2: Tất cả các cửa sổ chứa khuôn mặt từ tầng 1 sẽ được sàng lọc bằng cách đưa vào một CNN khác gọi là Mạng lọc (R-Net) để tiếp tục loại bỏ một số lượng lớn các cửa sổ không chứa khuôn mặt. Sau đó, thực hiện hiệu chuẩn với véc-tơ hồi quy và thực hiện hợp nhất các cửa sổ xếp chồng nhau tại một vùng.



Hình 4. *Mạng lọc (R-Net) [5]*

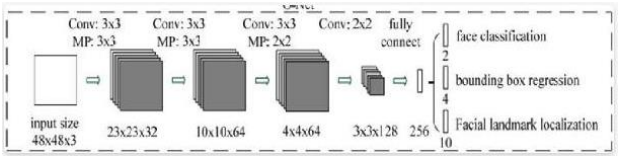
Trong bước R-Net sử dụng kiến trúc CNN gồm: 3 lớp tích chập, 2 lớp co và 1 lớp kết nối đầy đủ. Đầu vào cửa sổ trượt với kích thước 24x24x3 (3 tương ứng với 3 màu: Đỏ, xanh lục, xanh lam trong hệ màu RGB thông thường). Kết quả của R-Net phân được 3 cụm:

- Cụm thứ nhất có 2 lớp nhận dạng khuôn mặt;

- Cụm thứ hai có 4 lớp đánh dấu vị trí hộp giới hạn;

- Cụm thứ ba có 10 lớp vị trí khuôn mặt.

Tầng 3: Tầng này tương tự như tầng 2, sử dụng CNN chi tiết nhất được gọi là Mạng đầu ra (O-Net) để lọc kết quả một lần nữa và đánh dấu vị trí năm điểm chính trên khuôn mặt.



Hình 5. *Mạng đầu ra (O-Net) [5]*

Mạng O-Net sử dụng CNN gồm: 4 lớp tích chập, 2 lớp co, 1 lớp kết nối đầy đủ. Đầu vào cửa sổ trượt có kích thước 48x48x3 (trong đó số 3 tương ứng với 3 màu: Đỏ, xanh lục, xanh lam trong hệ màu RGB thông thường). Kết quả của O-Net phân được 3 cụm:

- Cụm thứ nhất có 2 lớp nhận dạng khuôn mặt;
- Cụm thứ hai có 4 lớp đánh dấu vị trí hộp giới hạn;
- Cụm thứ ba có 10 lớp vị trí khuôn mặt.

Ứng dụng MTCNN để phát hiện khuôn mặt cho phép xác định khuôn mặt trong bức ảnh tốt hơn so với các phương pháp khác.

2.3. Thuật toán FaceNet nhận dạng khuôn mặt

Các thuật toán nhận dạng khuôn mặt trước đây chủ yếu biểu diễn khuôn mặt bằng một véc-tơ đặc trưng và thông qua một lớp bottleneck để giảm số chiều dữ liệu. Tuy nhiên, số chiều dữ liệu của véc-tơ đặc trưng thường tương đối lớn nên sẽ làm cho tốc độ nhận dạng giảm xuống. Vì vậy, thuật toán PCA thường được áp dụng để giảm số chiều dữ liệu của véc-tơ đặc trưng và tăng tốc độ nhận dạng. Đồng thời, trong các phương pháp nhận dạng thì hàm loss function thường chỉ xác định khoảng cách giữa 2 bức ảnh (đại lượng mô tả sự giống nhau của hai bức ảnh). Như vậy, xuất hiện vấn đề là trong một lần huấn luyện chỉ có thể học được một kết quả: Hoặc là giống nhau nếu hai bức ảnh cùng thuộc về một lớp, hoặc là khác nhau nếu hai bức ảnh thuộc về hai lớp riêng.

FaceNet là một thuật toán hỗ trợ cho việc nhận dạng và phân cụm khuôn mặt cho phép giải quyết các hạn chế nêu trên [6]. FaceNet sử dụng một mạng CNN và cho phép giảm số chiều dữ liệu của véc-tơ đặc trưng (thường sử dụng là 128 chiều). Do đó, cho phép tăng tốc độ huấn luyện và xử lý mà độ chính xác vẫn được đảm bảo. Đối với thuật toán FaceNet, hàm loss function sử dụng hàm triplet loss cho phép khắc phục hạn chế của các phương pháp nhận dạng trước đây, quá trình huấn luyện cho phép học được đồng thời: Sự giống nhau giữa hai bức ảnh (nếu hai bức ảnh cùng một lớp) và sự khác nhau giữa hai bức ảnh (nếu chúng không cùng một lớp).

FaceNet chính là một dạng siam network thường biểu diễn véc-tơ đặc trưng của các bức ảnh trong một không gian Euclidean n chiều (thường là 128 chiều). Việc biểu diễn thường tuân theo quy tắc: Nếu khoảng cách giữa các véc-tơ embedding càng nhỏ, thì mức độ tương đồng giữa chúng càng lớn và ngược lại. Tập hợp véc-tơ này sẽ là dữ liệu đầu vào cho hàm loss function để đánh giá chỉ số khoảng cách giữa các véc-tơ.

FaceNet sử dụng CNN bằng cách dùng hàm f(x) và nhúng hình ảnh x vào không gian Euclidean d chiều sao cho khoảng cách giữa các hình ảnh của 1 người không phụ thuộc vào điều kiện bên ngoài, khoảng cách giữa các khuôn mặt giống nhau (của cùng một người là nhỏ) trong khi khoảng cách giữa các ảnh khác nhau sẽ có khoảng cách lớn.

Hàm $f(x) \in \mathbb{R}^d$ có chức năng biểu diễn ảnh x vào không gian Euclidean d chiều. Tại đây, sẽ có 3 bức ảnh là: Anchor (ảnh gốc), Positive (ảnh gần giống với ảnh gốc) và Negative (ảnh khác với ảnh gốc). Sau khi biểu diễn vào không gian Euclide thì tương ứng với 3 bức ảnh trên là X_a, X_p và X_n. Để nhận dạng tốt thì khoảng cách từ X_a tới X_p sẽ phải nhỏ hơn khoảng cách từ X_a tới X_n:

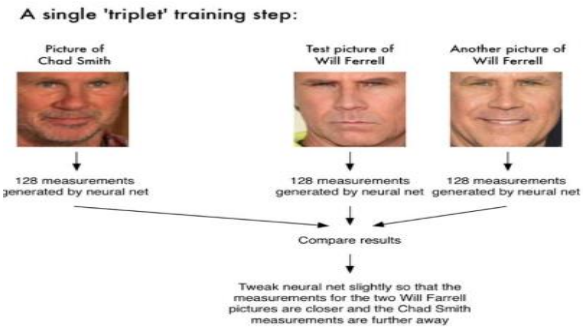
$d(X_a, X_p) < d(X_a, X_n)$

Do đó, dẫn đến biểu thức 1 (với g là giá trị biên):

$$||f(Xa) - f(Xp)||_2^2 + g < ||f(Xa) - f(Xn)||_2^2 \quad (1)$$

Lúc đó hàm triplet loss sẽ có dạng như sau:

$$L(Xa, Xp, Xn) = \sum ||f(Xa) - f(Xp)||_2^2 - ||f(Xa) - f(Xn)||_2^2 + g \quad (2)$$



Hình 6. *Minh hoạ bộ ba sai số*

Khi huấn luyện mô hình siam network với triplet loss cần phải xác định trước cặp ảnh (X_a, X_p) thuộc về cùng một người. Ảnh X_n là ảnh khác với ảnh gốc của người đó thường sẽ được lựa chọn ngẫu nhiên từ các bức ảnh thuộc các lớp còn lại. Do đó, tập hợp ảnh X_n thường được thu thập nhiều hơn 1 bức ảnh/1 người để có thể chuẩn bị được tập dữ liệu huấn luyện. Nếu 1 người chỉ có 1 ảnh thì có thể đưa những tập dữ liệu như vậy làm bộ ảnh X_n khi huấn luyện.

Như đã nêu trên có thể thấy, khi sử dụng triplet loss vào các mô hình CNN có thể tạo ra các véc-tơ đặc trưng tốt nhất cho mỗi một bức ảnh. Các véc-tơ đặc trưng này sẽ cho phép phân biệt rõ các ảnh Negative (ảnh khác với ảnh gốc) rất giống ảnh Positive (ảnh gần giống với ảnh gốc). Hơn nữa, khoảng cách giữa các bức ảnh thuộc cùng một lớp sẽ trở nên gần nhau hơn trong không gian chiếu Euclidean.

Tuy vậy, việc sử dụng bộ ba như trên sẽ khiến cho quá trình hội tụ chậm. Do đó, cần chọn bộ ba thích hợp trong quá trình huấn luyện để cải thiện được hiệu suất và độ chính xác của mô hình.

Để khắc phục được việc hội tụ chậm, thường sẽ chọn bộ ba sai số sao cho khoảng cách giữa ảnh gốc và ảnh gần với ảnh gốc (ảnh của cùng 1 người) là lớn nhất và khoảng cách giữa ảnh gốc và ảnh của người khác là gần nhất:

$$argmax(||f(X_a) - f(X_p)||_2^2)$$

$$argmax(||f(X_a) - f(X_n)||_2^2)$$

Việc chọn hình ảnh như trên có thể xảy ra trường hợp

$$||f(X_a) - f(X_p)||_2^2 > ||f(X_a) - f(X_n)||_2^2 \quad (3)$$

Lúc này ta sẽ huấn luyện làm sao cho biểu thức (3) trở về biểu thức (2). Việc huấn luyện sẽ giúp khoảng cách giữa hai ảnh của cùng 1 người là nhỏ nhất và ngược lại ảnh của 2 người sẽ có khoảng cách là lớn nhất.

Hình 7. *Minh họa về quá trình sau huấn luyện [6]*

Việc lựa chọn bộ ba sai số sẽ ảnh hưởng đến hiệu quả của mô hình, nếu giá trị bộ ba sai số được xác định tốt thì quá trình hội tụ khi huấn luyện sẽ nhanh hơn và kết quả sẽ cho độ chính xác cao hơn. Việc lựa chọn ngẫu nhiên bộ ba sai số có thể dẫn tới mô hình huấn luyện không thể hội tụ.

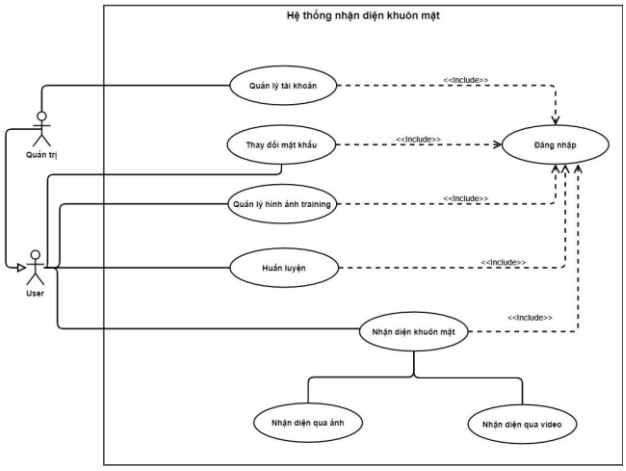
3. Xây dựng hệ thống và triển khai đánh giá

3.1. Xây dựng hệ thống nhận dạng khuôn mặt

Nhóm tác giả tiến hành xây dựng hệ thống nhận dạng trên cơ sở ứng dụng MTCCN và FaceNet sử dụng mạng nơ-ron tích chập và thuật toán softmax.

Hệ thống gồm 02 phân quyền chính là Quản trị viên và Người dùng. Phân quyền Người dùng có thể thực hiện các chức năng: Đăng nhập, quản lý tập hình ảnh huấn luyện, huấn luyện mô hình nhận dạng và nhận dạng khuôn mặt (nhận dạng thông qua hình ảnh hoặc nhận dạng trực tiếp: Sử dụng camera). Phân quyền quản trị thừa kế từ phân quyền người dùng và có thêm chức năng quản lý tài khoản người dùng.

Chức năng hệ thống được mô tả thông qua biểu đồ sau:



Hình 8. *Biểu đồ ca sử dụng tổng quan của hệ thống*

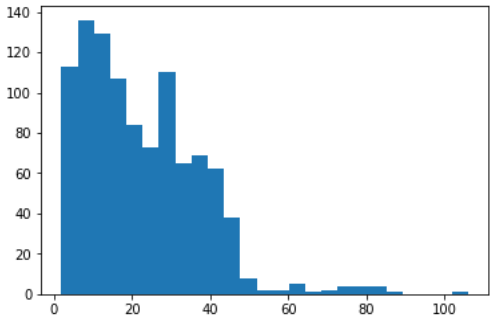
Bên cạnh đó, backend của hệ thống được xây dựng dựa trên API (sử dụng thư viện flask), sử dụng thư viện imgur để lưu trữ hình ảnh, Tensorflow để thực hiện các phép tính toán và sử dụng colaboratory của google để huấn luyện dữ liệu.

3.2. Triển khai thử nghiệm và đánh giá hệ thống

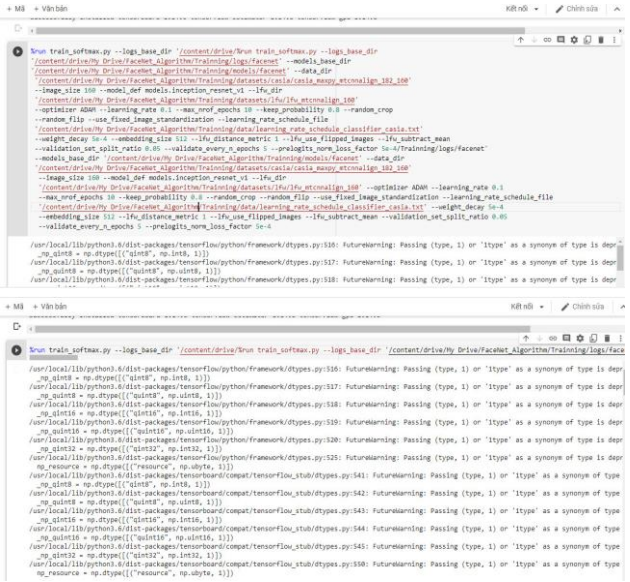
Để huấn luyện mô hình, nhóm tác giả sử dụng tập dữ liệu khuôn mặt người Việt từ Google Image (gồm 23105 khuôn mặt của 1020 người).

Đặc điểm của bộ dữ liệu này các ảnh của một người được thu thập tại các thời kỳ, hoàn cảnh khác nhau.

Nhóm tác giả lấy ảnh của 980 người để làm tập dữ liệu huấn luyện mô hình. Mô hình được huấn luyện với tham số như Hình 10 và thời gian huấn luyện mô hình rơi vào khoảng 60 phút.



Hình 9. *Số lượng hình ảnh phân bố của tập dữ liệu khuôn mặt người Việt*



Hình 10. *Tham số huấn luyện mô hình*

Tập dữ liệu để kiểm tra bao gồm 40 người với 874 bức ảnh. Nhóm tác giả sử dụng 574 ảnh để huấn luyện và 300 ảnh để kiểm thử (tất cả các ảnh được điều chỉnh về kích thước 160x160) cho cả 2 thuật toán: Eigengaces-PCA (sử dụng haar cascade để phát hiện khuôn mặt) và FaceNet (sử dụng MTCNN để phát hiện khuôn mặt). Kết quả thực hiện nhận dạng được thể hiện ở Bảng 1.

Dữ liệu thực tế từ Bảng 1 cho thấy, thời gian nhận dạng trung bình của phương pháp Eigengaces-PCA nhanh hơn với phương pháp đề xuất. Tuy nhiên, phương pháp nhận dạng khuôn mặt sử dụng thuật toán FaceNet và MTCNN để phát hiện khuôn mặt cho kết quả nhận dạng chính xác cao hơn.

Bảng 1. *Kết quả nhận dạng khuôn mặt sử dụng Eigengaces-PCA và FaceNet*

Phương pháp	Số ảnh huấn luyện	Số ảnh kiểm tra	Số ảnh nhận diện đúng	Số ảnh nhận diện sai	Thời gian nhận dạng trung bình (giây)	Hiệu suất
Eigengaces-PCA (haar cascade)	574	300	262	38	0,27	87,33%
FaceNet (MTCNN)	574	300	285	15	0,43	95%

4. Kết luận

Bài báo nghiên cứu xây dựng hệ thống nhận dạng khuôn mặt trên cơ sở áp dụng MTCNN và thuật toán FaceNet (sử dụng không gian Euclidean) để phát hiện và nhận dạng khuôn mặt, cho phép cải thiện độ chính xác khi nhận dạng. Kết quả thực nghiệm cho thấy, hệ thống có thể áp dụng đối với các bài toán nhận dạng khuôn mặt trong thực tế.

TÀI LIỆU THAM KHẢO

[1] Shaimaa Khudhair Salah, Waleed Rasheed Humood, Ahmed Othman Khalaf, "A Proposed Generalized Eigenfaces System for Face Recognition Based on One Training Image", *Journal of Southwest Jiaotong University*, Volume 55, No 2, 2020, pp. 1-11.

[2] Frank Peprah, Michael Asante, "Comparative Analysis Of The Performance Of Principal Component Analysis (PCA) And Linear Discriminant Analysis (LDA) As Face Recognition Techniques", *International Journal of Scientific & Technology Research*, Volume 6, Issue 10, 2017, pp. 286-291.

[3] R. Shyam, Y.N. Singh, "Face recognition using augmented local binary patterns and bray curtis dissimilarity metric", *Proc. 2nd Int. Conf. Signal Processing and Integrated Network (SPIN 2015). IEEE; (2015)*, 2015.

[4] José Augusto Cadena Moreano, Nora Bertha La Serna Palomino, "Global Facial Recognition Using Gabor Wavelet, Support Vector Machines and 3D Face Models", *Journal of Advances in Information Technology*, Vol. 11, No. 3, 2020, pp. 143-148.

[5] Kaipeng Zhang, Zhanpeng Zhang, Zhifeng Li, Yu Qiao, "Joint Face Detection and Alignment using Multi-task Cascaded Convolutional Networks", *IEEE Signal Processing Letters*, Volume: 23, Issue: 10, 2016, pp. 1499-1503.

[6] Florian Schroff, Dmitry Kalenichenko, James Philbin, "FaceNet: A Unified Embedding for Face Recognition and Clustering", *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, 2015.